

520.43091X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): INABA, et al
Serial No.: Not yet assigned
Filed: September 2, 2003
Title: METHOD, APPARATUS AND PROGRAMS FOR DELIVERING
INFORMATION
Group: Not yet assigned

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

September 2, 2003

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2002-256802, filed September 2, 2002.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl L. Brundidge
Registration No. 29,621

CIB/alb
Attachment
(703) 312-6600

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 2 年 9 月 2 日
Date of Application:

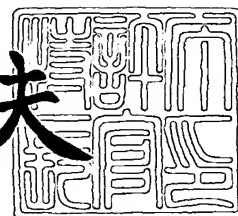
出 願 番 号 特 願 2 0 0 2 - 2 5 6 8 0 2
Application Number:
[ST. 10/C]: [J P 2 0 0 2 - 2 5 6 8 0 2]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 3 年 8 月 1 3 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 3 - 3 0 6 5 3 6 8

【書類名】 特許願

【整理番号】 NT02P0519

【提出日】 平成14年 9月 2日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 17/30

【発明者】

【住所又は居所】 神奈川県川崎市幸区鹿島田 8 9 0 番地 株式会社日立製作所 ビジネスソリューション事業部内

【氏名】 稲場 靖彦

【発明者】

【住所又は居所】 神奈川県川崎市幸区鹿島田 8 9 0 番地 株式会社日立製作所 ビジネスソリューション事業部内

【氏名】 松林 忠孝

【発明者】

【住所又は居所】 神奈川県川崎市幸区鹿島田 8 9 0 番地 株式会社日立製作所 ビジネスソリューション事業部内

【氏名】 弥生 隆明

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

【氏名】 内角 真

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報配信方法、その装置及びそのプログラム

【特許請求の範囲】

【請求項 1】

情報源から送られた文書情報に対してユーザにより設定された配信条件を適用し該配信条件を満たす文書情報を該ユーザに向けて配信する情報配信方法において、該ユーザに配信した文書情報を記憶手段に保存し、該ユーザから該配信条件の変更要求を受けたとき、保存された該文書情報に変更後の配信条件を適用し該変更後の配信条件を満たさないために該ユーザに配信されないことになる文書情報を該ユーザに提示することを特徴とする情報配信方法。

【請求項 2】

情報源から送られた文書情報に対してユーザについて設定されたプロファイルを適用して該プロファイルに基づく該文書情報の適合度を計算し、適合条件を満たす文書情報を該ユーザに向けて配信する情報配信方法において、該プロファイルが変更されるとき変更前の該プロファイルを記憶手段に保存し、受け取った文書情報に対して変更後のプロファイルにより適合条件を満たす第 1 の文書情報を該ユーザに向けて配信し、該受け取った文書情報に対して該変更前のプロファイルにより適合条件を満たす文書情報のうち該第 1 の文書情報に含まれない文書情報を該ユーザに提示することを特徴とする情報配信方法。

【請求項 3】

情報源から送られた文書情報に対してユーザについて設定された適合プロファイル及び非適合プロファイルを適用して該文書情報の適合度を計算し、適合条件を満たす文書情報を該ユーザに向けて配信し、配信した該文書情報の少なくとも 1 つに対して該ユーザから不適の評価を受けたとき、該不適の評価を受けた文書情報に含まれる特徴文字列を該非適合プロファイルに追加するようにフィードバックする情報配信方法において、

該ユーザに配信した文書情報を記憶手段に保存し、該保存された該文書情報に該フィードバックした後のプロファイルを適用して適合度を計算し、その結果として適合条件を満たさないために該ユーザに配信されないことになる文書情報を

該ユーザに提示し、該ユーザの応否に応じて該フィードバックを有効とするか取り消すことを特徴とする情報配信方法。

【請求項 4】

情報源から送られた文書情報に対してユーザについて設定された適合プロフィール及び非適合プロフィールを適用して該文書情報の適合度を計算し、適合条件を満たす文書情報を該ユーザに向けて配信し、配信した該文書情報の少なくとも 1 つに対して該ユーザから不適の評価を受けたとき、該不適の評価を受けた文書情報に含まれる特徴文字列を該非適合プロフィールに追加するようにフィードバックする情報配信方法において、

該ユーザに配信した文書情報を記憶手段に保存し、該保存された該文書情報のうち該不適の評価を受けた文書情報との類似度が所定値より高い文書情報を該ユーザに提示し、該ユーザの応否に応じて該不適の評価を受けた文書情報と該類似度が所定値より高い文書情報に含まれる特徴文字列を該非適合プロフィールに追加するか該不適の評価を受けた文書情報に固有の特徴文字列を該非適合プロフィールに追加するフィードバックを行うことを特徴とする情報配信方法。

【請求項 5】

情報源から送られた文書情報に対してユーザにより設定された配信条件を適用し該配信条件を満たす文書情報を該ユーザに向けて配信する手段と、該ユーザに配信した文書情報を保存する記憶手段と、該ユーザから該配信条件の変更要求を受けたとき、保存された該文書情報に変更後の配信条件を適用し該変更後の配信条件を満たさないために該ユーザに配信されないことになる文書情報を該ユーザに提示する手段とを有することを特徴とする計算機。

【請求項 6】

情報源から送られた文書情報に対してユーザについて設定されたプロフィールを適用して該プロフィールに基づく該文書情報の適合度を計算する手段と、適合条件を満たす文書情報を該ユーザに向けて配信する手段と、該プロフィールが変更されるとき変更前の該プロフィールを保存する記憶手段と、受け取った文書情報に対して変更後のプロフィールにより適合条件を満たす第 1 の文書情報を該ユーザに向けて配信する手段と、該受け取った文書情報に対して該変更前のプロフ

ファイルにより適合条件を満たす文書情報のうち該第 1 の文書情報に含まれない文書情報を該ユーザに提示する手段とを有することを特徴とする計算機。

【請求項 7】

コンピュータに、情報源から送られた文書情報に対してユーザにより設定された配信条件を適用し該配信条件を満たす文書情報を該ユーザに向けて配信する機能、該ユーザに配信した文書情報を記憶手段に保存する機能、該ユーザから該配信条件の変更要求を受けたとき、保存された該文書情報に変更後の配信条件を適用し該変更後の配信条件を満たさないために該ユーザに配信されないことになる文書情報を該ユーザに提示する機能を実現させるためのプログラム。

【請求項 8】

コンピュータに、情報源から送られた文書情報に対してユーザについて設定されたプロファイルを適用して該プロファイルに基づく該文書情報の適合度を計算する機能、適合条件を満たす文書情報を該ユーザに向けて配信する機能、該プロファイルが変更されるとき変更前の該プロファイルを記憶手段に保存する機能、受け取った文書情報に対して変更後のプロファイルにより適合条件を満たす第 1 の文書情報を該ユーザに向けて配信する機能、および該受け取った文書情報に対して該変更前のプロファイルにより適合条件を満たす文書情報のうち該第 1 の文書情報に含まれない文書情報を該ユーザに提示する機能を実現させるためのプログラム。

【請求項 9】

コンピュータに、情報源から送られた文書情報に対してユーザについて設定された適合プロファイル及び非適合プロファイルを適用して該文書情報の適合度を計算する機能、適合条件を満たす文書情報を該ユーザに向けて配信する機能、配信した該文書情報の少なくとも 1 つに対して該ユーザから不適の評価を受けたとき、該不適の評価を受けた文書情報に含まれる特徴文字列を該非適合プロファイルに追加するようにフィードバックする機能、該ユーザに配信した文書情報を記憶手段に保存する機能、該保存された該文書情報に該フィードバックした後のプロファイルを適用して適合度を計算する機能、その結果として適合条件を満たさないために該ユーザに配信されないことになる文書情報を該ユーザに提示する機

能、および該ユーザの応否に応じて該フィードバックを有効とするか取り消す機能を実現させるためのプログラム。

【請求項 10】

コンピュータに、情報源から送られた文書情報に対してユーザについて設定された適合プロファイル及び非適合プロファイルを適用して該文書情報の適合度を計算する機能、適合条件を満たす文書情報を該ユーザに向けて配信する機能、配信した該文書情報の少なくとも 1 つに対して該ユーザから不適の評価を受けたとき、該不適の評価を受けた文書情報に含まれる特徴文字列を該非適合プロファイルに追加するようにフィードバックする機能、該ユーザに配信した文書情報を記憶手段に保存する機能、該保存された該文書情報のうち該不適の評価を受けた文書情報との類似度が所定値より高い文書情報を該ユーザに提示する機能、および該ユーザの応否に応じて該不適の評価を受けた文書情報と該類似度が所定値より高い文書情報に含まれる特徴文字列を該非適合プロファイルに追加するか該不適の評価を受けた文書情報に固有の特徴文字列を該非適合プロファイルに追加するフィードバックを行う機能を実現させるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ユーザが予め登録しておいた配信条件に合致する文書情報を各ユーザに配信する情報配信技術に関する。

【0002】

【従来の技術】

【特許文献 1】 特開 2000-339346 号公報

【特許文献 2】 特開 2001-256253 号公報

【特許文献 3】 特開平 11-143902 号公報

近年、電子メール等によりニュースなどの大量の電子化文書（以下、文書情報と呼ぶ）が時々刻々ユーザへ配信されるようになってきた。また WWW（World Wide Web）を利用して情報発信を行う情報源も急増しており、これらの情報源から情報収集ロボット等を用いて収集される文書情報も膨大な量とな

っている。このため、これらの文書情報の中からユーザが求める情報を含む文書情報を検索し、これを該ユーザに配信する情報フィルタリングシステムへのニーズが高まっている。

【0003】

この情報フィルタリングシステムの一例として、特開2000-339346号公報（特許文献1）が開示されている。特許文献1は、ユーザが配信して欲しい情報を表すサンプルの文書（以下、種文書と呼ぶ）を配信条件として予め入力する。そして文書情報が発生した際には、情報フィルタリングシステムが各ユーザの種文書と各文書情報との内容の適合度を所定の方法により算出し、適合度が所定の閾値を超える文書情報だけを該当ユーザに配送する。

【0004】

またさらに受け取った配信結果をユーザが評価し、その評価を反映して以降の配信精度を向上するレリバンスフィードバックと呼ばれる技術がある。レリバンスフィードバックにおいては、ユーザが受け取った配信結果文書それぞれに対して、「欲しい情報である」という評価（以下、適と評価、と呼ぶ）や「欲しい情報ではなかった」（以下、不適と評価、と呼ぶ）を入力し、それをもとにユーザの配信条件を保持したデータ（以下、プロファイルと呼ぶ）を修正する。この技術を利用した情報フィルタリングシステムの一例として特開2001-256253号公報（特許文献2）が開示されている。

【0005】

特許文献2を用いた情報フィルタリングシステムの処理の概要を図2に示す。特許文献2では、まずユーザ201が入力した種文書202から、その種文書の内容を特徴的に表す文字列（以下、特徴文字列と呼ぶ）を抽出し、適合プロファイル203に登録する（①）。本図に示した例では、「メジャーリーグ」（アメリカのプロの野球リーグ）に関する情報を欲しいユーザ201が「メジャーリーグの話題」という種文書202を設定している。ここで特徴文字列の抽出方法としては、特開平11-143902号公報（特許文献3）が開示されている方法と同一の方法でも良いし、形態素解析等を用いて文書中に出現する単語を抽出する方法や、単純n-gramを抽出する方法であってもかまわない。その後、フ

フィルタリングモジュール 206 が情報リソース 205 から発生した文書情報 207 を受け取ったとき、適合プロファイル 203 と各文書情報 207 との適合度を算出する。ここで、適合度算出式としては、例えば以下のようなものを用いる。

【0006】

【数 1】

$$S(D) = \sum_i^N \{ \text{Frq}(i) \times w(i) \} \quad \dots (\text{数 } 1)$$

この式で、 $S(D)$ は文書情報 D とプロファイルとの適合度であり、 $\text{Frq}(i)$ は特徴文字列 i の文書 D における出現回数であり、 $w(i)$ は該当プロファイル内の特徴文字列 i の重みである。 Σ は該当プロファイル内の全ての特徴文字列について和をとることを示す。この式により、プロファイル内で高い重みを付与された特徴文字列が多く出現する文書情報ほど、高い適合度が算出されることになる。この適合度が所定の閾値を超えた文書情報を当該ユーザ 201 に配信する。図 2 に示した例では、文書情報 207 のうち所定の閾値を超えた文書情報 208、209、210 のみをユーザ 201 に送信する (②)。

【0007】

ここで、これらのうち文書情報 210 は、「メジャーリーグ」の話題ではなく「サッカー」の話題であったため、ユーザ 201 にとって欲しい情報ではなかったとする。このとき特許文献 2 では、ユーザ 201 から「文書情報 210 は欲しい情報ではない」との入力を受け、文書情報 210 から特徴文字列を抽出し、ユーザが欲しくない概念を表すデータを保持した非適合プロファイル 204 に、それらの特徴文字列を追加する (③)。

【0008】

そしてこれ以降、情報フィルタリングシステム 206 は、情報リソース 205 から送信されてきた文書情報のうち、適合プロファイル 203 との適合度が所定の閾値より高く、かつ非適合プロファイル 204 との適合度が所定の閾値より低い文書情報のみを選別してユーザに送信する。以上示した特許文献 2 の仕組みにより、配信された文書情報に対するユーザの評価を利用して、対話的に配信する文書情報の精度を高めていくことが可能になる。

【0009】

【発明が解決しようとする課題】

しかしながら、特許文献2のような機能を備えた情報フィルタリングシステムにおいては、以下に示すような問題がある。文書には複数の観点が含まれる場合がある。例えば「サッカー」の話題が中心の文書であっても、その一部分に「メジャーリーグ」の話題が記述されている場合がある。「サッカー」の話題は不要と考えるユーザがこの文書は「欲しい情報ではない」と評価してしまうと、「メジャーリーグ」に関する話題もユーザが所望していないと判断してプロフィールを修正してしまう可能性がある。そのため、ユーザの意図に反して欲しい情報が送られなくなってしまう場合がある。加えて、ユーザが文書情報の内容を読み違えてしまったり、操作を誤るなどして、間違った評価を入力してしまった場合、それ以降ユーザには欲しい情報が送られてこなくなってしまう可能性がある。

【0010】

例えば図2に示した例において、「メジャーリーグ」の話題が欲しいユーザ201が「欲しい情報ではない」と評価した文書情報210の一部分に「メジャーリーグ」に関する話題が記述されていた場合を考える。このとき、文書情報210から「メジャー」、「リーグ」などといったユーザ201が欲しい概念を表す特徴文字列が抽出され、非適合プロフィール204に追加されてしまう可能性がある。したがって、これ以降ユーザ201が欲しい「メジャーリーグ」に関する情報が送られなくなる可能性がある。

【0011】

このように欲しい情報が送られてこなくなった場合、ユーザは本当に欲しい情報が発生していないのか、それとも自分の評価が意図と反して解釈されたためにプロフィールが不適切になってしまったのか、判断ができないという問題がある。すなわちユーザは、ユーザの意図に反するフィードバックがなかったとしたら取得できたはずの文書情報を手に入れることが不可能である上に、欲しい情報が配信されない理由も判断できないという問題がある。

【0012】

以上のような問題を解決するため、本発明の目的は、ユーザが配信条件を変更

するような操作をしたとき、それによって配信されないことになる文書情報をユーザに提示し、配信条件の変更を評価可能とすることにある。

【0 0 1 3】

【課題を解決するための手段】

本発明は、情報源から送られた文書情報に対してユーザにより設定された配信条件を適用しこの配信条件を満たす文書情報をそのユーザに向けて配信する情報配信方法において、ユーザに配信した文書情報を記憶手段に保存し、そのユーザから配信条件の変更要求を受けたとき、保存された文書情報に変更後の配信条件を適用し、変更後の配信条件を満たさないためにユーザに配信されないことになる文書情報をそのユーザに提示する情報配信技術を特徴とする。

【0 0 1 4】

【発明の実施の形態】

以下、本発明の実施の形態について説明する。

【0 0 1 5】

まず本発明の第一の実施例について説明する。第一の実施例は、対話的なインタフェースにより、ユーザの意図をより正しく反映したレリバンスフィードバックにユーザを誘導し、意図に合わないレリバンスフィードバックが行われてしまうことを防止するものである。

【0 0 1 6】

本実施例による情報フィルタリングシステムのシステム構成を図 1 に示す。本実施例における情報フィルタリングシステム 1 0 は、ディスプレイ 1 0 0、キーボード 1 0 1、中央演算処理装置（C P U） 1 0 2、主メモリ 1 0 4 およびこれらを結ぶバス 1 0 3 から構成される。またバス 1 0 3 には、L A N（L o c a l A r e a N e t w o r k）等の通信回線 1 0 5 を介して、文書情報を配信する文書情報配信元 1 0 6 と情報フィルタリングシステム 1 0 を利用するユーザ 1 0 7 が接続されている。情報配信元 1 0 6 及びユーザ 1 0 7 は、計算機あるいは情報フィルタリングシステム 1 0 に接続する端末装置である。文書情報配信元 1 0 6 は、電子メール等を用いて電子化された文書情報を本システム 1 0 へ配信したり、インターネットを介して文書情報を提示する。文書情報配信元 1 0 6 として

は、通信社や新聞社等の情報発生源などが考えられるが、その他のものであっても構わない。ユーザ107は電子メールを用いて配信条件を本システム10へ登録する。本システムからはその配信条件に基づいて検索された文書情報が電子メールを用いて該当ユーザへ配信される。

【0017】

以下本発明の実施例では、文書情報配信元106は電子メール等を用いて本システム10に文書情報を配信するものとして述べるが、文書情報配信元106がインターネット上に提示した文書情報を、情報収集ロボットを用いて収集するものとしてもかまわない。またユーザ107は、電子メールを用いて配信条件を本システム10へ登録するものとしても良いし、インターネットを介して本システムへ登録するものとしても良い。さらに以下に詳述する配信条件に基づいて、文書情報が本システム10から電子メールを用いて該当ユーザに配信されるものとして述べるが、インターネット等を介して提示するようにしてもかまわない。

【0018】

主メモリ104には、メインプログラム108、プロファイル生成プログラム109、フィルタリングプログラム110、適合フィードバックプログラム112、非適合フィードバックプログラム113、適合プロファイル116、および非適合プロファイル114が格納される。これらプログラムは、CPU102によって実行される。以上のプログラムやユーザプロファイルは、ハードディスク装置（本図には示していない）、フレキシブルディスク（本図には示していない）等のコンピュータで読み書きできる記憶媒体に格納することもできる。

【0019】

メインプログラム108は、情報フィルタリングシステム10の管理者によるキーボード101からの指示を受け起動し、情報フィルタリングシステム10のシステム制御を行うプログラムである。処理の流れの詳細は後述する。

【0020】

プロファイル生成プログラム109は、ユーザ107から入力された種文書から特徴文字列を生成し、適合プロファイル116に格納するプログラムである。ここでプロファイルの生成方法及び生成される適合プロファイルの内容は、図2

で説明した通りである。ここで適合プロファイル 116 は、ユーザ 107 が欲しい概念を表す特徴文字列を格納したものであり、非適合プロファイル 114 はユーザ 107 が欲しくない概念を表す特徴文字列を格納したものである。

【0021】

フィルタリングプログラム 110 は、文書情報配信元 106 から受信した文書情報を、その文書情報を欲しいと思っているユーザに送信するプログラムである。本プログラムは、例えば特開 2000-339346 号公報（特許文献 1）に開示されている技術を用い、適合プロファイル 116 および非適合プロファイル 114 を適用して情報配信元 106 から受信した各文書情報との適合度を所定の方法により算出する。その適合度が所定の閾値より高い場合、ユーザ 107 の配信条件を満たしたとして当該文書情報をユーザ 107 に配信するプログラムである。

【0022】

適合フィードバックプログラム 112 および非適合フィードバックプログラム 113 は、ユーザ 107 が本システム 10 から受信した各文書情報に対する「適」または「不適」といった評価にもとづき、各々該当ユーザの適合プロファイル 116 または非適合プロファイル 114 を修正するプログラムである。処理の流れの詳細は後述する。

【0023】

以下、第一の実施例におけるメインプログラム 108 の処理手順について図 3 の PAD (Problem Analysis Diagram) 図を用いて説明する。メインプログラム 108 は、まずステップ 301 においてステップ 302 からステップ 309 までのステップを、システム管理者によってシステムを終了されるまで繰り返す。まずステップ 302 で、ユーザ 107 から種文書が入力されたと判定した場合、ステップ 303 においてプロファイル生成プログラム 109 を起動し、該ユーザ 107 の適合プロファイル 116 を設定する。

【0024】

次にステップ 304 で情報リソース 106 から文書情報が送られてきたと判定した場合、ステップ 305 においてフィルタリングプログラム 110 を起動し、

各ユーザ107のプロファイルと該文書情報との適合度を算出し、その適合度が所定の閾値を越えているユーザに該文書情報を配信する。適合度の算出方法の例として、文書情報に出現する特徴文字列のうち適合プロファイル116に登録されている特徴文字列についてその重みを適用して（数1）により第1の適合度を計算し、非適合プロファイル114に登録されている特徴文字列についてその重みを適用して（数1）により第2の適合度を計算し、第1の適合度から第2の適合度を引くことによって目的の適合度を求める方法がある。

【0025】

次にステップ306で、配信した文書情報の少なくとも1つに対する適という評価がユーザ107から入力されたと判定した場合、ステップ307で適合フィードバックプログラム112を起動し、該ユーザ107の適合プロファイル116を修正する。次にステップ308で、配信した文書情報の少なくとも1つに対する不適という評価がユーザ107から入力されたと判定した場合、ステップ309で非適合フィードバックプログラム113を起動し、該ユーザ107の非適合プロファイル114を修正する。適合フィードバックプログラム112および非適合フィードバックプログラム113によるプロファイル修正方法については後述する。以上がメインプログラム108の処理手順である。

【0026】

次にステップ307および309においてそれぞれ起動される適合フィードバックプログラム112及び非適合フィードバックプログラム113について説明する。これらのプログラムは、ユーザが受け取った文書情報に対し「適」や「不適」という評価を入力した場合に、その評価によってユーザの意図にそぐわないフィードバックが行なわれる可能性を判定し、可能性が高い場合にユーザに確認を行い、ユーザの意図に合ったフィードバックを実現するプログラムである。ここではまずユーザが「不適」の評価を入力した場合に起動される非適合フィードバックプログラム113の処理手順について図4のPAD図を用いて説明する。

【0027】

非適合フィードバックプログラム113は、まずステップ401において、ユーザ107が不適と評価した文書情報から特徴文字列を抽出する。次にステップ

402において、該ユーザ107にこれまでに配送した記憶装置に保存されている文書情報（図示せず）のうち不適と評価していないものについて、ステップ403から405の処理を繰り返す。まずステップ403において、該当する文書の1つとユーザが不適と評価した文書との類似度を算出する。ここで類似度の算出方法としては、ベクトル空間法を用いても良いし、これら文書から特徴文字列を抽出し類似度算出用の一時的なプロファイルを作成したうえで（数1）を用いて算出するなど、その他の方法を用いても良い。ユーザが不適と評価した文書と類似度の高い過去の文書は、不適の疑いのある文書ということになる。次にステップ404において、ステップ403で算出した類似度が所定の類似度よりも高いと判定する場合、ステップ405で該当する文書から特徴文字列を抽出する。

【0028】

次にステップ406において、ステップ404の判定条件を満たす文書がひとつ以上存在する場合、ステップ407からステップ411を実行する。まずステップ407において、ステップ401とステップ407で抽出した特徴文字列をもとに、確認画面を表示する。

【0029】

ここでこの確認画面の内容について、図5を用いて説明する。ここでは、ユーザが「メジャーリーグに続きW杯サッカーが開幕」というタイトルの文書情報に対して不適と評価した場合を例にとる。まず過去にそのユーザに配信した文書情報のうち、ユーザが不適と評価した文書情報502と内容が類似している文書情報503を提示する。ここで文書同士の類似度はステップ403で算出される。また文書情報502および文書情報503に出現する特徴文字列のうち、ユーザが不適と評価した文書情報502にのみ出現する特徴文字列504と、文書情報503に出現する特徴文字列505を表示する。ユーザはこの情報を参照して、特徴文字列504だけでなく特徴文字列505もユーザが欲しくない概念を表すものであると判定した場合は「はい」ボタン506を押下し、そうでない場合は「いいえ」ボタン507を押下する。また、レリバンスフィードバック処理をキャンセルしたい場合は、「キャンセル」ボタン508を押下する。

【0030】

次にステップ408において、ステップ407で表示した確認画面においてユーザ107が「はい」ボタン506を押下した場合、ステップ409において、ステップ401とステップ405で抽出した特徴文字列を該当ユーザの非適合プロフィール114に追加する。一方ステップ410において、ステップ407で表示した確認画面においてユーザ107が「いいえ」ボタン507を押下した場合、ステップ411において、ステップ401で抽出した特徴文字列のうち、ステップ405で抽出されなかったもののみを該当ユーザの非適合プロフィール114に追加する。

【0031】

ステップ409およびステップ411で抽出される特徴文字列は、ユーザにとって望ましくない文字列とともに望ましい文字列が混入するおそれがある。抽出される特徴文字列を適合プロフィール116に登録される文字列と照合し、すでに適合プロフィール116に登録されている特徴文字列を除き残りの特徴文字列を非適合プロフィール114に追加すれば、このようなノイズの文字列の混入を防止できる。

【0032】

また適合フィードバックプログラム112の処理手順は、図4に示す非適合フィードバックプログラム113の処理手順とほぼ同様である。ただしステップ409とステップ411において特徴文字列を追加する先は適合プロフィール116となる点が異なる。またステップ409およびステップ411で抽出される特徴文字列は、ユーザにとって望ましい文字列とともに望ましくない文字列が混入するおそれがある。抽出される特徴文字列を非適合プロフィール114に登録される文字列と照合し、すでに非適合プロフィール114に登録されている特徴文字列を除き残りの特徴文字列を適合プロフィール116に追加すれば、このようなノイズの文字列の混入を防止できる。

【0033】

以上示した適合フィードバックプログラム112および非適合フィードバックプログラム113により、以下のことが可能となる。例えばユーザがある文書情報に不適という評価を入力した場合、図5で説明した画面においてユーザに提示

された特徴文字列 504 および 505 が共にユーザの欲しくない概念を表すものである場合は、それらを共に非適合プロファイル 114 に追加することができる。すなわち、ユーザが欲しくない概念を表す特徴文字列を、漏らすことなく非適合プロファイル 114 に追加することができる。

【0034】

一方、提示した特徴文字列 504 および 505 のうち、ユーザが不適と評価した文書に出現する特徴文字列 504 のみがユーザの欲しくない概念を表すものである場合は、ユーザが不適と評価した文書にのみ出現する特徴文字列 504 のみを非適合プロファイル 114 に追加することができる。すなわち、ユーザが本当は欲しい概念を表す特徴文字列を非適合プロファイル 114 に追加してしまうことを防止できる。

【0035】

またユーザにとっての別の判断方法として、文書情報 503 はいまままでの配信条件では配信していたが、レリバンスフィードバックによって今後配信されなくなり得る文書情報のサンプルと考えることができる。したがってユーザは、文書情報 503 のような文書情報が今後配信されないと困ると感じた場合には、「いいえ」ボタン 507 を押下することにより、ユーザが本当は今後も欲しい情報を取得することが可能である。以上より、ユーザの意図に合わないレリバンスフィードバックが行なわれてしまうことを防止し、ユーザの意図をより正しく反映したレリバンスフィードバックが可能となる。

【0036】

次に本発明の第二の実施例について説明する。第二の実施例は、ユーザがこれから変更しようとする配信条件が自分にとって適切なものであるか否かを判断可能とするものである。これにより、ユーザの意図に合わない配信条件に変更され欲しい文書情報が取得できなくなってしまうことを防止するものである。

【0037】

本実施例による情報フィルタリングシステムのシステム構成を図 6 に示す。本実施例によるシステム構成は、第一の実施例と同様の形態をとるが、以下の点において異なる。主メモリ 104 には、第一の実施例におけるプロファイル生成プ

プログラム 109、適合フィードバックプログラム 112、非適合フィードバックプログラム 113 の代わりに、配信条件設定プログラム 601、文書情報保存プログラム 602、条件変更リハーサルプログラム 603 が格納される。またメインプログラム 606 は、第一の実施例のものと処理の内容が異なる。また主メモリ 104 には、配信情報格納エリア 605 及び文書情報格納エリア 604 が確保される。

【0038】

配信条件設定プログラム 601 は、ユーザが入力した配信条件を配信条件格納エリア 605 に格納する。ここで配信条件とは、キーワードやキーワードのブーリアン条件（論理和条件や論理積条件など）で記述されるものを想定するが、文書情報配信元の識別子や配信日付の範囲指定条件など、その他の配信条件でも良い。文書情報保存プログラム 602 は、各ユーザ 107 に配信した文書情報を、どのユーザに配信したかという情報とともに文書情報格納エリア 604 に保存するプログラムである。条件変更リハーサルプログラム 603 は、ユーザ 107 から配信条件格納エリア 605 に格納されている当該ユーザの配信条件の変更要求が入力された場合に、その変更の結果、配信結果がどう変わるかを当該ユーザに例示するプログラムである。

【0039】

以下、第二の実施例におけるメインプログラム 606 の処理手順について図 7 の PAD 図を用いて説明する。メインプログラム 606 は、まずステップ 701 においてシステム管理者によってシステム 10 が終了されるまで、以下のステップ 702 からステップ 708 の処理を繰り返す。まずステップ 702 で、ユーザ 107 から新規の配信条件が入力されたと判定した場合は、ステップ 703 において配信条件設定プログラム 601 を起動し、当該ユーザの配信条件を設定し、配信条件格納エリア 605 に書き込む。

【0040】

次にステップ 704 で文書情報配信元 106 から文書情報が送られてきたと判定した場合、ステップ 705 においてフィルタリングプログラム 110 を起動し、配信条件格納エリア 605 中の各ユーザの配信条件を満たしているかを判定し

、配信条件を満たしているユーザに該文書情報を配信する。また続くステップ706において、文書情報保存プログラム602を起動し、各ユーザに配信した文書情報を文書情報格納エリア604に保存する。

【0041】

次にステップ707でユーザから配信条件変更要求が入力されたと判定した場合には、ステップ708において条件変更リハーサルプログラム603を起動し、当該ユーザに配信条件変更の確認画面を表示する。条件変更リハーサルプログラム603の処理の流れ及びこのとき表示される画面の詳細については後述する。以上が、メインプログラム606の説明である。

【004.2】

次に図7に示したステップ708においてメインプログラム606により起動される条件変更リハーサルプログラム603の処理手順について、図8のPAD図を用いて説明する。本プログラムは前述したとおり、ユーザ107から配信条件格納エリア605に格納されている当該ユーザの配信条件の変更要求が入力された場合に、その変更の結果、配信結果がどう変わるかを当該ユーザに例示するプログラムである。

【0043】

まずステップ1201において、文書情報格納エリア604に格納されている文書情報のうち該ユーザに過去に配信した文書を対象に、新たに設定するよう要求された配信条件により検索を実行する。次にステップ1202において、ステップ1201で検索にヒットしなかった文書情報を当該ユーザに提示し、警告する。ここで当該ユーザに提示される画面の例を図9に示す。

【0044】

この確認画面901では、まず当該ユーザが新たに設定しようとする配信条件902を表示する。そして当該ユーザに過去に配信した文書情報のうち、配信条件902を満たさない文書情報903を提示する。このことにより、新しく設定しようとしている配信条件902によると配信されなくなってしまう文書情報の例がわかる。したがって、配信条件902がユーザにとって適切か否か判断することができ、適切であると判断した場合は「はい」ボタン904を、適切でない

と判断した場合は「いいえ」ボタン 905 を押下する。

【0045】

次にステップ 1203 において、ステップ 1202 で提示した確認画面で、ユーザが「はい」ボタン 904 を押下した場合、ステップ 1204 において配信条件格納エリア 605 内の当該ユーザの配信条件を変更要求のとおり変更する。次にステップ 1205 において、ステップ 1202 で提示した確認画面で、ユーザが「いいえ」ボタン 905 を押下した場合、ステップ 1206 において配信条件の変更を中止する。

【0046】

以上説明したように、条件変更リハーサルプログラム 603 により、ユーザ 107 はこれから変更しようとする配信条件が、自分にとって適切なものであるかを判断することができ、誤った配信条件に変更してしまうことにより、欲しい文書情報が取得できなくなってしまうことを防止できる。なお提示される文書情報 903 がすべてユーザにとって不要であれば、「はい」ボタン 904 の押下でよい。文書情報 903 に不要なものが必要なものが混在している場合には、「いいえ」ボタン 905 を押下し、配信条件 902 を修正し、再び提示される文書情報 903 をレビューすればよい。

【0047】

なお第一の実施例で示したレリバンスフィードバックを備えた情報配信システムについて、レリバンスフィードバックしたプロファイルが適当か否かをみるために、過去にユーザに配信した文書情報でテストするという手順をとってもよい。この第一の実施例の変形システムは、第一の実施例のプロファイル生成プログラム 109、フィルタリングプログラム 110、適合フィードバックプログラム 112、非適合フィードバックプログラム 113、適合プロファイル 116、および非適合プロファイル 114 と、第二の実施例の文書情報保存プログラム 602、条件変更リハーサルプログラム 603、文書情報格納エリア 604 とを備える。ただし条件変更リハーサルプログラム 603 は、処理内容が変更となる。適合フィードバックプログラム 112 又は非適合フィードバックプログラム 113 が新しい特徴文字列を各々適合プロファイル 116 又は非適合プロファイル 11

4に追加した後、本例の条件変更リハーサルプログラム603は、文書情報格納エリア604に格納されている文書情報のうち当該ユーザに配信したものを対象に、ステップ305の方法によって対象文書の適合度を算出し、ヒットしなかったもの、すなわち今後配信されなくなるものをユーザに提示し、警告する。条件変更リハーサルプログラム603は、ユーザが「はい」の指示をした場合には、レリバンスフィードバックを有効とみなし、ユーザが「いいえ」の指示をした場合には、フィードバック取り消しの処理をする。このことにより、レリバンスフィードバックによって今後の配信結果がどう変わるかを知ることができるため、この時点でユーザがレリバンスフィードバックの適否を判断できるし、プロファイルの修正を中止することも可能となる。したがって、ユーザの意図に合わないフィードバックにより、欲しい文書情報が取得できなくなってしまうことを防止できる。

【0048】

以下、本発明の第三の実施例を説明する。第三の実施例は、ユーザの意図に合わないレリバンスフィードバックが行なわれた場合でも、ユーザが欲しい文書情報を手に入れることを可能とし、かつレリバンスフィードバックを取り消してプロファイルを以前のものに戻すことを可能とするものである。

【0049】

本実施例による情報フィルタリングシステムのシステム構成を図10に示す。本実施例によるシステム構成は、第一の実施例と同様の形態をとるが、以下の点において異なる。主メモリ104には、第一の実施例のものと同一プロファイル生成プログラム109、フィルタリングプログラム110、適合フィードバックプログラム112、非適合フィードバックプログラム113のほかに、追加情報表示プログラム1001、およびフィードバック取り消しプログラム1002が格納される。また第一の実施例のものと同一適合プロファイル116および非適合プロファイル114のほかに、変更前ユーザプロファイル1003が格納される。またメインプログラム1010の処理内容が、第一の実施例におけるメインプログラム108とは異なる。

【0050】

追加情報表示プログラム 1001 は、適合プロファイル 116 および非適合プロファイル 114 にもとづいて各ユーザ 107 に配信する文書情報のほかに、後述する追加条件を満たす文書情報を追加情報としてユーザに配信するプログラムである。処理の流れの詳細は後述する。フィードバック取り消しプログラム 1002 は、直前に行なったレリバンスフィードバックを行う前のプロファイルに戻すプログラムである。処理の流れの詳細は後述する。変更前プロファイル保存エリア 1003 は、各ユーザが一番最近に行なったレリバンスフィードバックを行う前の適合プロファイルと非適合プロファイル、すなわち最新のものよりひとつ古い適合プロファイルと非適合プロファイルを保存しておくエリアである。

【0051】

以下、第三の実施例におけるメインプログラム 1010 の処理手順について図 11 の PAD 図を用いて説明する。メインプログラム 1010 は、まずステップ 1101 においてシステム管理者によってシステム 10 が終了されるまで、以下のステップ 1102 からステップ 1110 の処理を繰り返す。まずステップ 1102 でユーザ 107 から種文書が入力されたと判定した場合、ステップ 1103 においてプロファイル生成プログラム 109 を起動し、このユーザ 107 の適合プロファイル 116 を生成する。

【0052】

次にステップ 1104 で情報リソース 106 から文書情報が送られてきたと判定した場合、ステップ 1105 とステップ 1106 の処理を行う。まずステップ 1105 においてフィルタリングプログラム 110 を起動し、各ユーザ 107 のプロファイルと該文書情報との適合度を算出し、その適合度が所定の閾値を越えているユーザに該文書情報を配信する。ステップ 1105 の処理はステップ 305 と同じ処理となる。続いてステップ 1106 で、追加情報表示プログラム 1001 を起動し、追加情報を当該ユーザに提示する。ここで提示する追加情報の詳細は後述する。

【0053】

このあと実行されるステップ 1107 からステップ 1110 は、第一の実施例のものと同一である。すなわちステップ 1105 及びステップ 1106 で配信し

た文書情報に対するユーザの適という評価が入力されたと判定した場合、ステップ1108で適合フィードバックプログラム112を起動し、該ユーザ107の適合プロファイル116を修正する。次にステップ1109で、配信した文書情報に対するユーザの不適という評価が入力されたと判定した場合、ステップ1110で非適合フィードバックプログラム113を起動し、該ユーザ107の非適合プロファイル114を修正する。

【0054】

次に図11で示したステップ1106においてメインプログラム1010によって起動される追加情報表示プログラム1001の処理の手順について、図12のPAD図を用いて説明する。追加情報表示プログラム1001は、まずステップ801において、各ユーザ107の現在の適合プロファイル116と文書情報配信元106から送られてきた文書情報との適合度を算出する。このときステップ1105とは異なり、適合プロファイル116のみを用い、非適合プロファイル114を適合度算出には用いない。次にステップ802において、ステップ801で算出した適合度が所定の閾値より高く、かつステップ1105で当該文書を配信していないユーザに、当該文書情報を配信する。このことにより、ユーザの適合プロファイルが反映されるが、非適合プロファイルが反映されないため、第一の実施例による方法では配信されない文書情報がユーザに配信されることになる。これにより、ユーザは自分の所望する概念を表す特徴文字列が含まれている文書情報を漏れなく受け取ることができる。またこれらは追加情報として提示されるため、ユーザは非適合プロファイルが反映されていないという事実を頭に入れて文書情報を参照することができる。そのため時間がないときには該情報は見ない、といった判断が可能である。

【0055】

次にステップ803において、変更前プロファイル保存エリア1003に保存されている各ユーザ107の現在よりひとつ前の適合プロファイルおよび非適合プロファイルを用いて、文書情報との適合度を算出する。適合度の算出手順は、ステップ1105と同じでよい。次にステップ804において、ステップ803で算出した適合度が所定の閾値より高く、かつステップ1105およびステップ

802において当該文書を配信していないユーザに、当該文書情報を配信する。このことにより、ユーザは一番最近のレリバンスフィードバックを行わなかった場合の配信結果を受け取ることができる。それゆえ、最新の配信条件とひとつ前の配信条件による配信結果を見比べることができ、一番最近のレリバンスフィードバックが適切なものであったか否かを判断することができる。かつレリバンスフィードバックが不適切な場合でも、それ以前の配信条件にもとづけば配信されるはずである所望の文書情報を取得することができる。

【0056】

ここまでのステップにより、ユーザに表示される画面の例を図13に示す。まずステップ1105により、従来の方法にもとづく配信文書情報1302が表示される。すなわち、ユーザの現在の適合プロファイル及び現在の非適合プロファイルの両方が反映された文書情報である。次にステップ802により、現在の適合度プロファイルが反映されるが非適合プロファイルが反映されない文書情報1303が提示される。そしてステップ804により、変更前のプロファイルと合致する文書情報1304が表示される。

【0057】

これらの文書情報をユーザが参照し、文書情報1302には欲しい情報が含まれておらず、文書情報1304に欲しい情報が含まれている場合には、前回行なったフィードバックがユーザの意図に合わないものであったと判断できる。その場合ユーザは「前回のフィードバックを取り消し」ボタン1305を押下する。また文書情報1303に欲しい情報が含まれているということは、現在の非適合プロファイル114が適切でない可能性がある。ユーザは、文書情報1303を参考にして非適合プロファイル114に含まれる特徴文字列の中で不要な特徴文字列を削除するかその重みを低減できる。

【0058】

次にステップ805において、ユーザから前回のフィードバックの取り消し要求が入力された（図13に示す表示画面において「前回のフィードバックを取り消し」ボタン1305を押下された）と判定した場合、ステップ806でフィードバック取り消しプログラム1002を起動し、変更前プロファイル保存エリア

1003に保存されている当該ユーザのひとつ前の適合プロファイルと非適合プロファイルを、該当ユーザの現在の適合プロファイル116および非適合プロファイル114にそれぞれ上書きする。

【0059】

このことにより、ユーザが最後に行ったレリバンスフィードバックが失敗だったと判断した場合、そのフィードバックを行う前の配信条件、すなわちプロファイルに戻すことが可能となる。

【0060】

なおここではユーザはひとつ前のプロファイルのみを保存するものとして説明したが、それ以上のものを保存しても良いし、全ての履歴を保存するものとしても良い。その場合、レリバンスフィードバックを何回でも遡ってやり直すことが可能となる。

【0061】

なお図13で説明した結果表示画面1301で表示される文書情報1302、1303、1304に対しても、適または不適の評価をユーザが入力して、レリバンスフィードバック処理を行うことができるようにしても良い。

【0062】

以上説明したように本実施例によれば、ユーザが過去に行なった文書情報に対する評価により、ユーザの意図に合わないレリバンスフィードバックが行なわれた場合でも、ユーザは欲しい文書情報を手に入れることが可能であり、かつレリバンスフィードバックを取り消してプロファイルを以前のものに返すことが可能となる。以上が本発明の第三の実施例の説明である。

【0063】

なお以上説明した実施例では、ユーザに配信する情報としては文書情報のみとしたが、画像や音声などその他のコンテンツを配信するものとしても良い。また以上の実施例において、ディスプレイ100、キーボード101、中央演算処理装置（CPU）102、主メモリ104、およびこれらを結ぶバス103から構成される情報フィルタリングシステムは、図1における文書情報配信元106と通信回線105の中間や、通信回線105とユーザ107の中間のいずれのネッ

トワーク上の位置に配置されても構わない。

【 0 0 6 4 】

【発明の効果】

以上述べたように本発明によれば、ユーザが配信条件を変更するような操作をしたとき、それによって配信されないことになる文書情報や特徴文字列をユーザに提示するので、ユーザは配信条件の変更を評価することができる。

【図面の簡単な説明】

【図 1】

本発明の第一の実施例による情報フィルタリングシステムのシステム構成図である。

【図 2】

従来技術を用いた情報フィルタリングシステムの処理の概要を示す図である。

【図 3】

第一の実施例におけるメインプログラム 1 0 8 の処理手順を示す P A D 図である。

【図 4】

第一の実施例における非適合フィードバックプログラム 1 1 3 の処理手順を示す P A D 図である。

【図 5】

第一の実施例において、ユーザが文書に評価を入力した際に表示される確認画面の例を示す図である。

【図 6】

本発明の第二の実施例による情報フィルタリングシステムのシステム構成図である。

【図 7】

第二の実施例におけるメインプログラム 6 0 6 の処理手順を示す P A D 図である。

【図 8】

第二の実施例における条件変更リハーサルプログラム 6 0 3 の処理手順を示す

PAD図である。

【図 9】

第二の実施例において、ユーザが配信条件の変更要求を入力した際に表示される確認画面の例を示す図である。

【図 10】

本発明の第三の実施例による情報フィルタリングシステムのシステム構成図である。

【図 11】

第三の実施例におけるメインプログラム 1010 の処理手順を示す PAD 図である。

【図 12】

第三の実施例における追加情報表示プログラム 1001 の処理の手順を示す PAD 図である。

【図 13】

第三の実施例において、ユーザに表示される配信情報表示画面の例を示す図である。

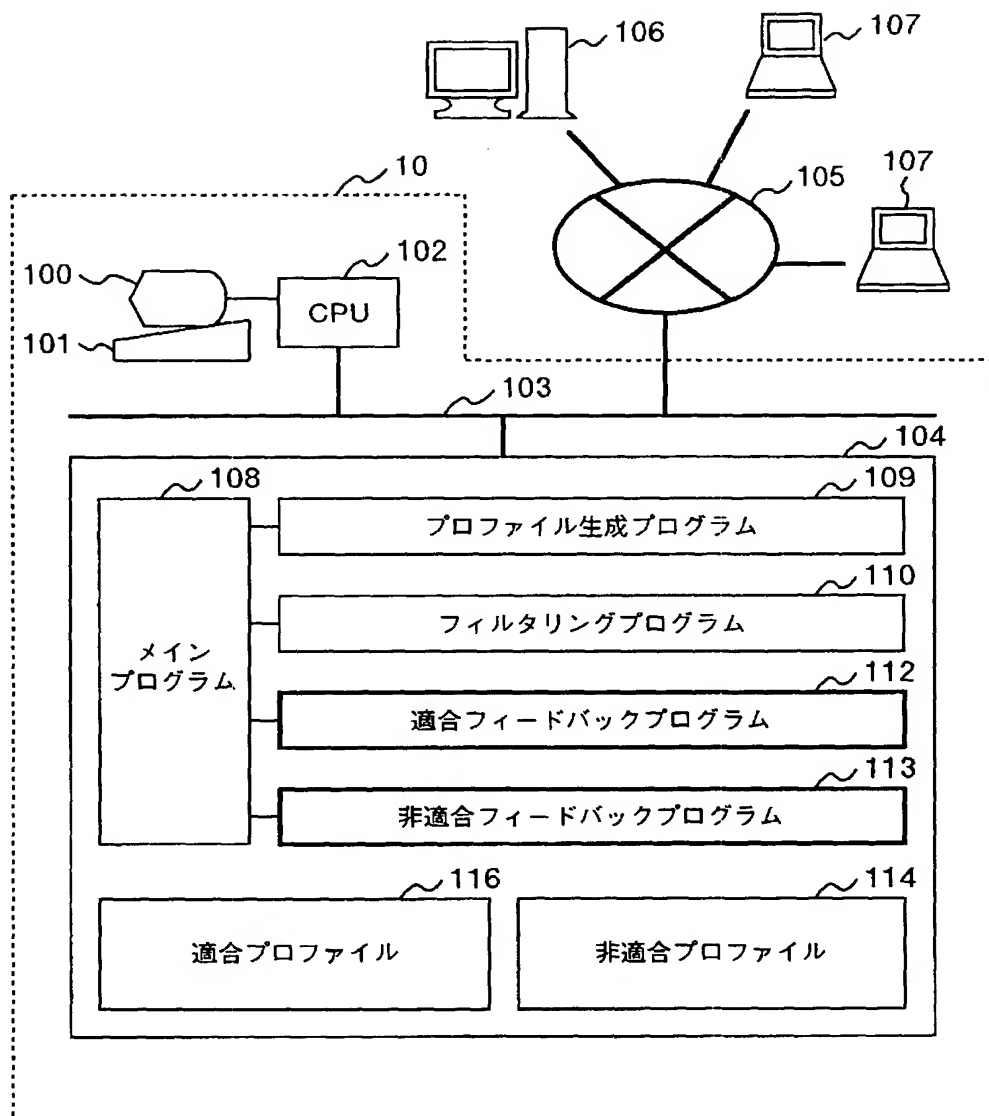
【符号の説明】

106：文書情報配信元、107：ユーザ、108、606、1010：メインプログラム、109：プロファイル生成プログラム、110：フィルタリングプログラム、112：適合フィードバックプログラム、113：非適合フィードバックプログラム、114：非適合プロファイル、116：適合プロファイル、601：配信条件設定プログラム、602：文書情報保存プログラム、603：条件変更リハーサルプログラム、604：文書情報格納エリア、605：配信条件格納エリア、1001：追加情報表示プログラム、1002：フィードバック取り消しプログラム、1003：変更前プロファイル保存エリア

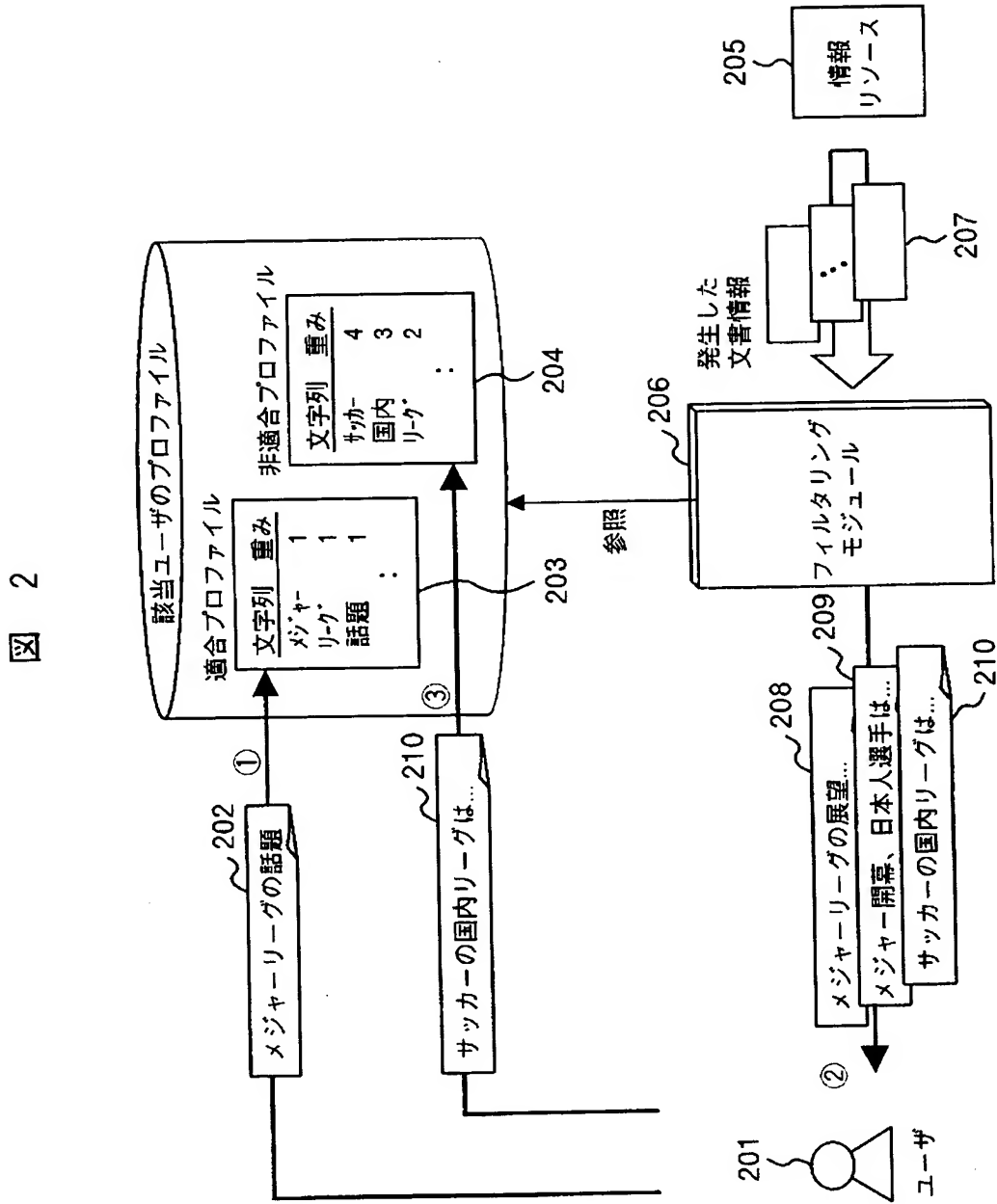
【書類名】 図面

【図 1】

図 1

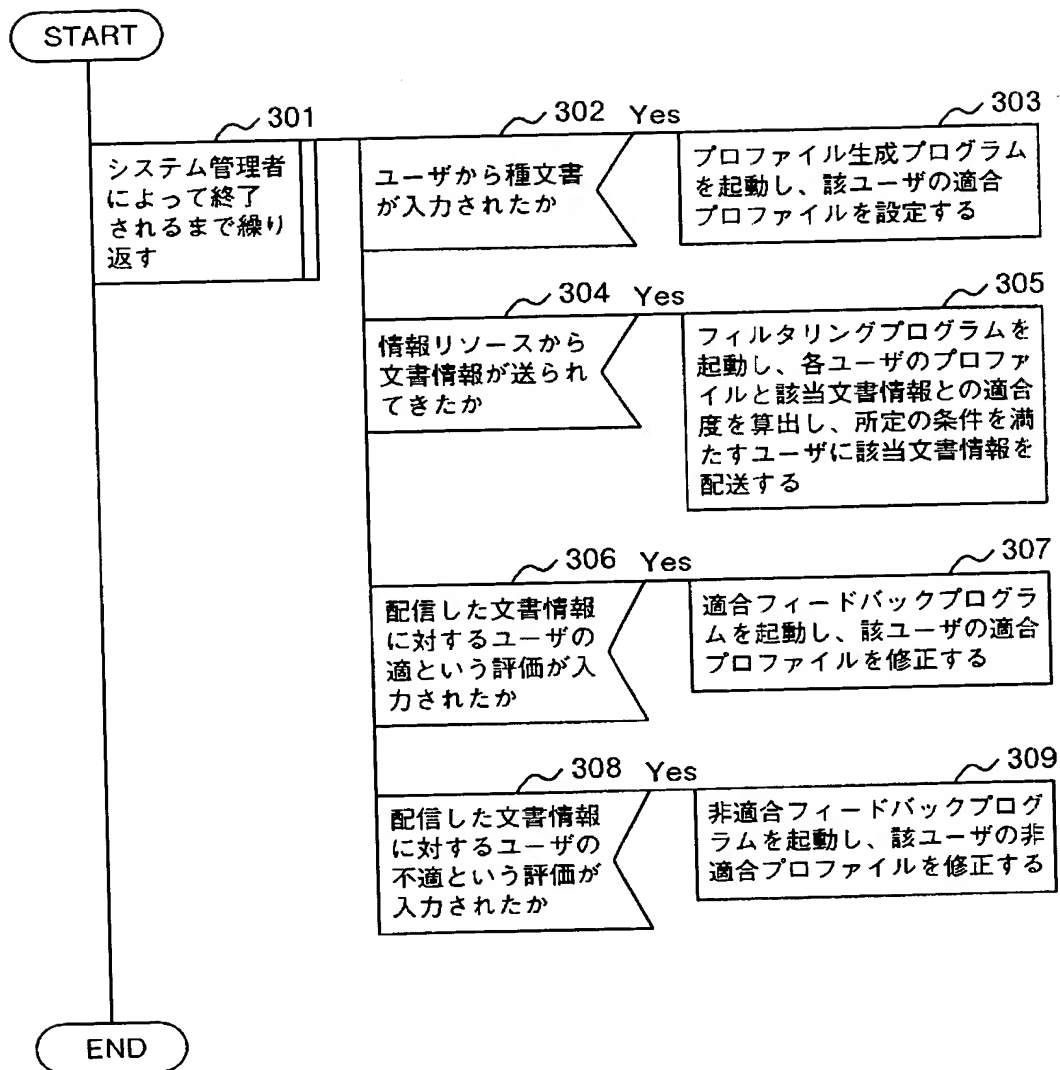


【図 2】



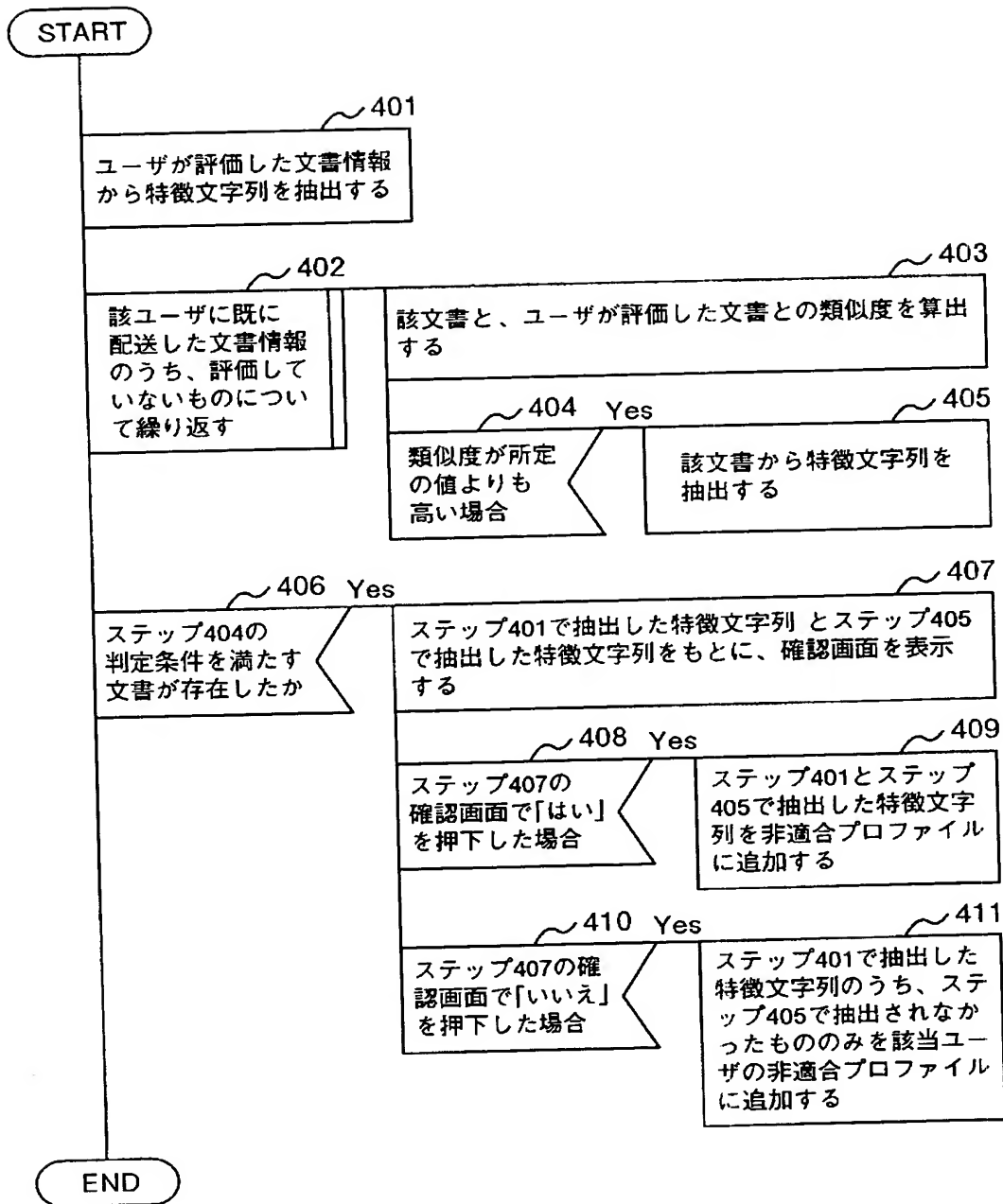
【図 3】

図 3



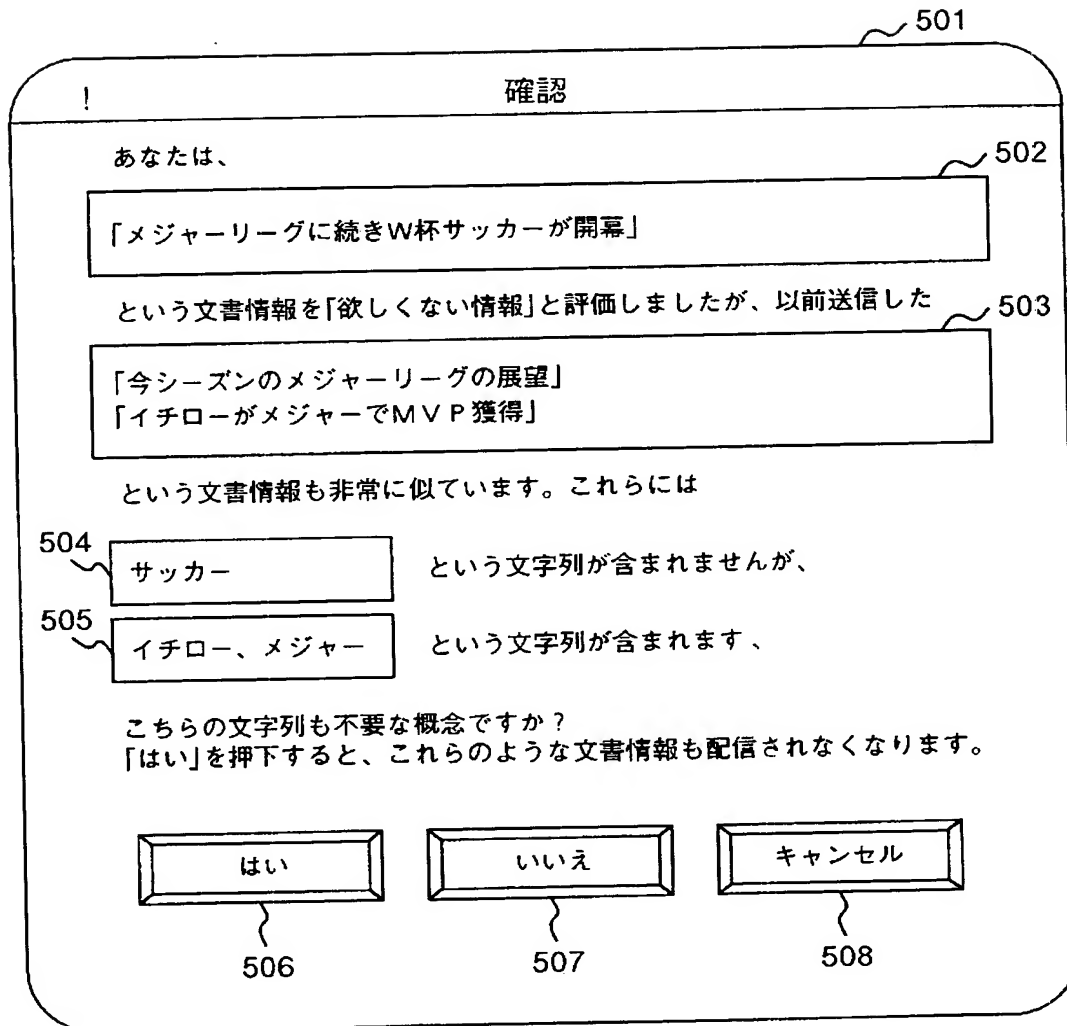
【図 4】

図 4



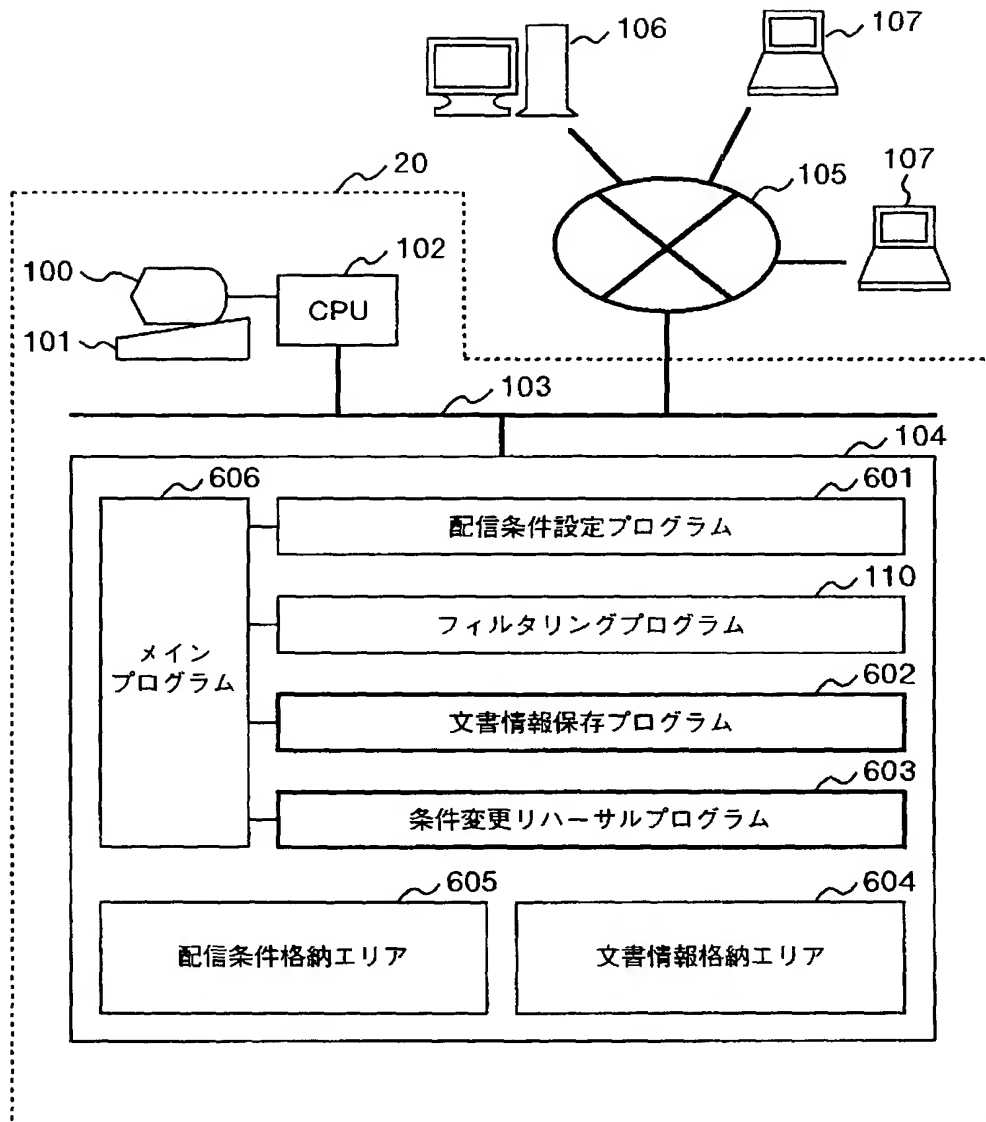
【図 5】

図 5



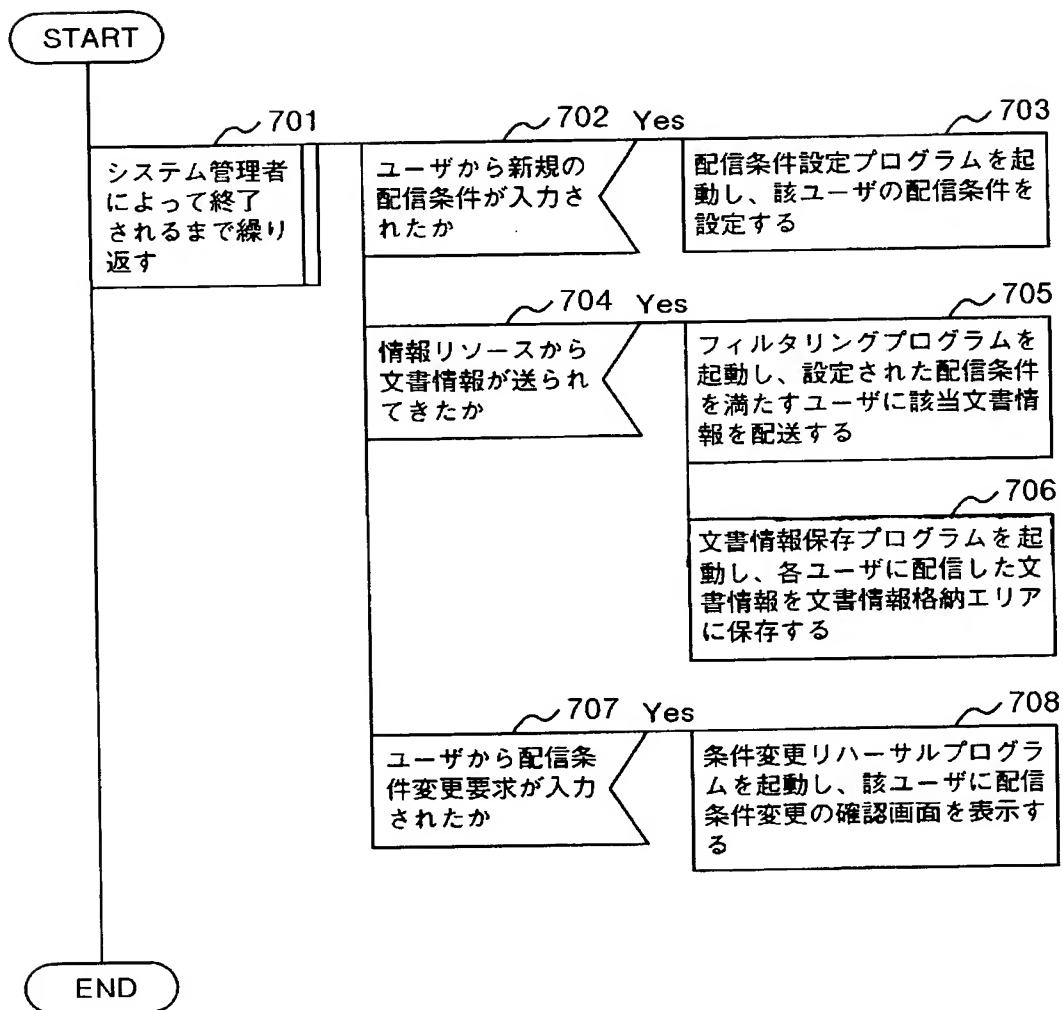
【図 6】

図 6



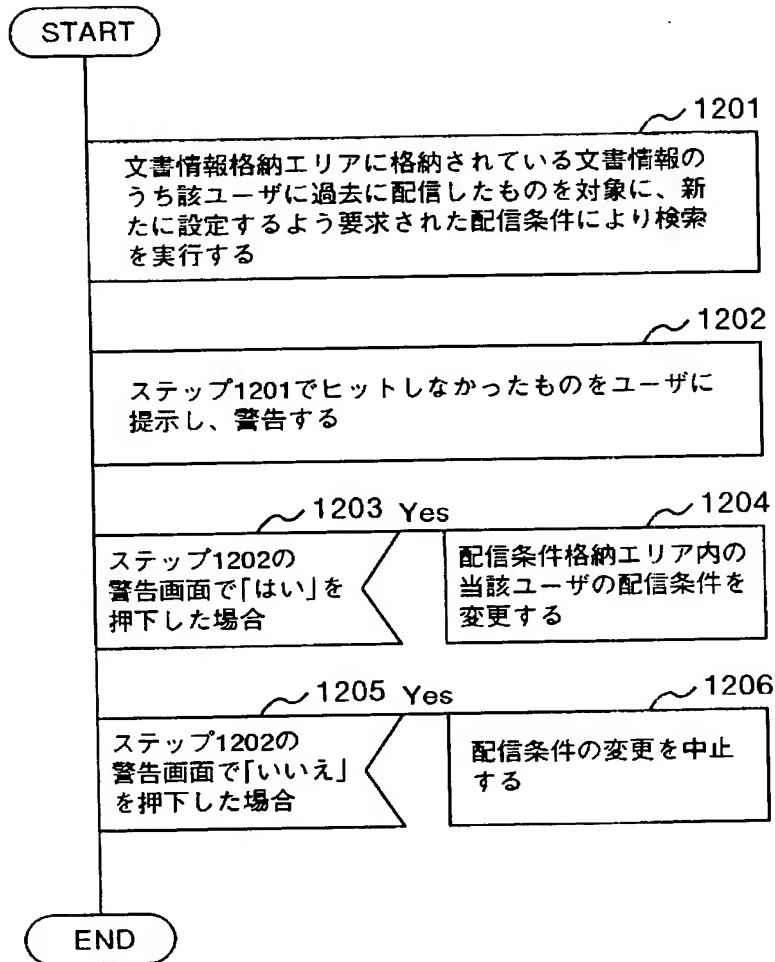
【図 7】

図 7



【図 8】

図 8



【図 9】

図 9

！ 確認

あなたは、

メジャー AND リーグ AND 野球

という配信条件に変更しましたが、この条件によると既に配信した

「今シーズンのメジャーリーグの展望」
「イチローがメジャーで MVP 獲得」

のような文書情報は今後配信されなくなります。

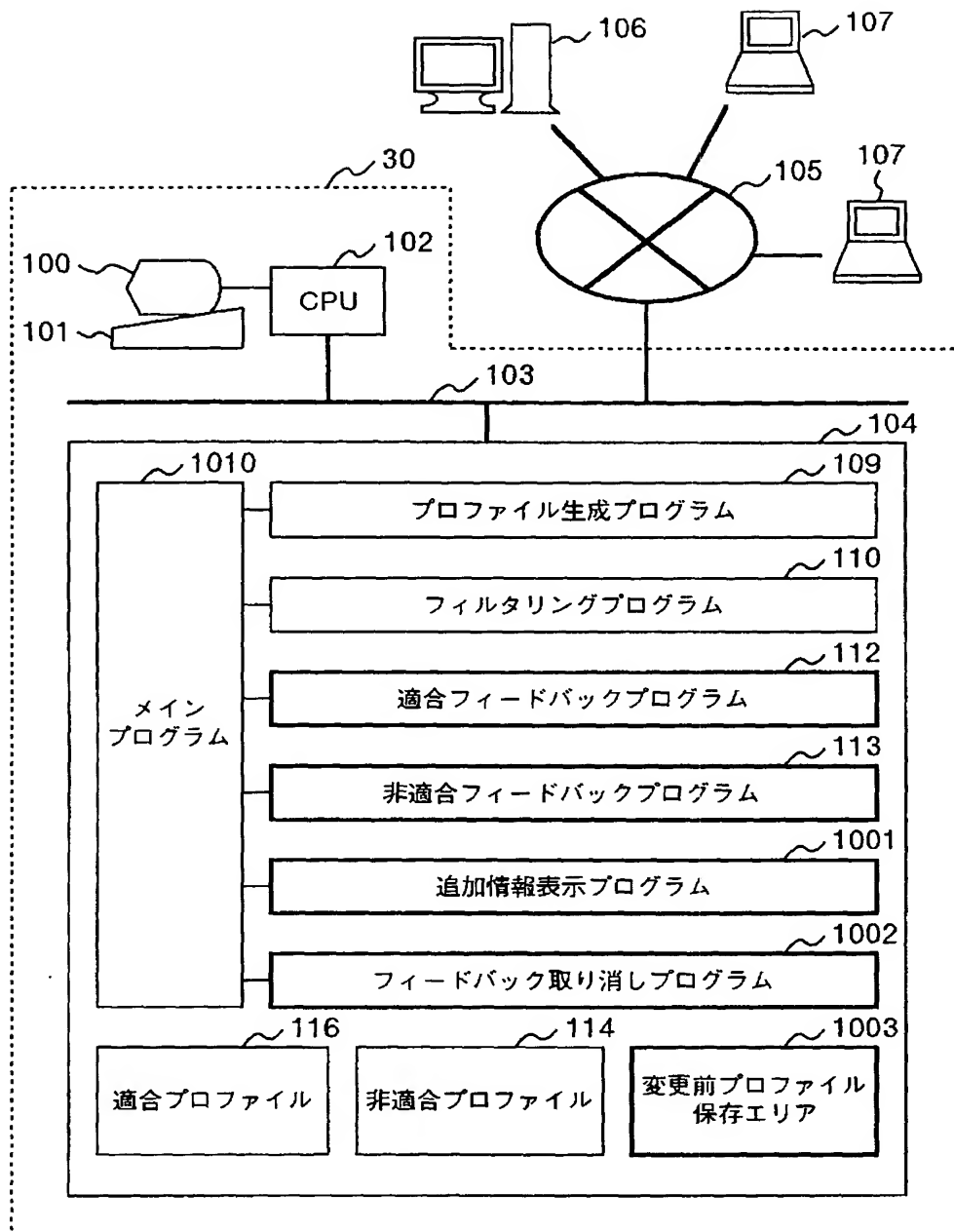
変更を確定してよろしいですか？

はい いいえ

904 905

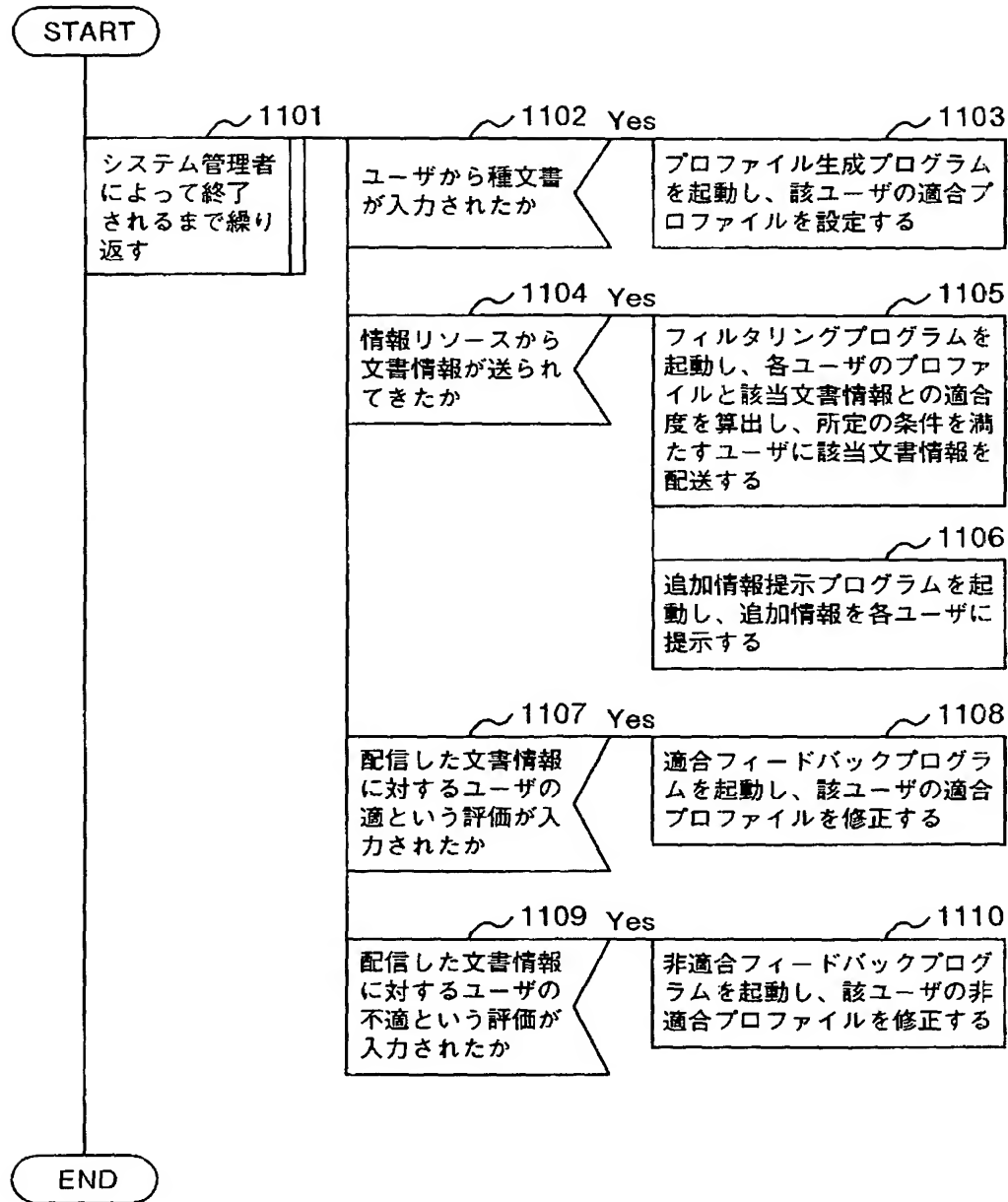
【図 10】

図 10



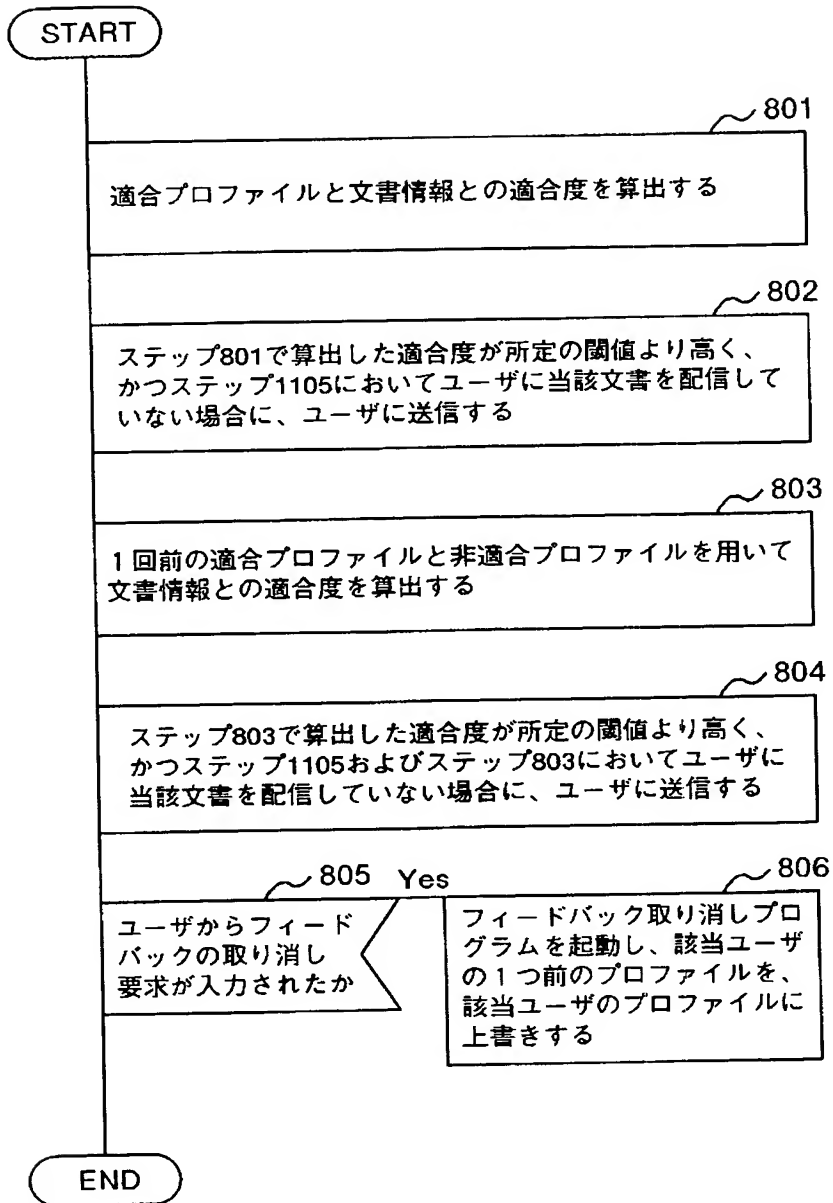
【図 11】

図 11



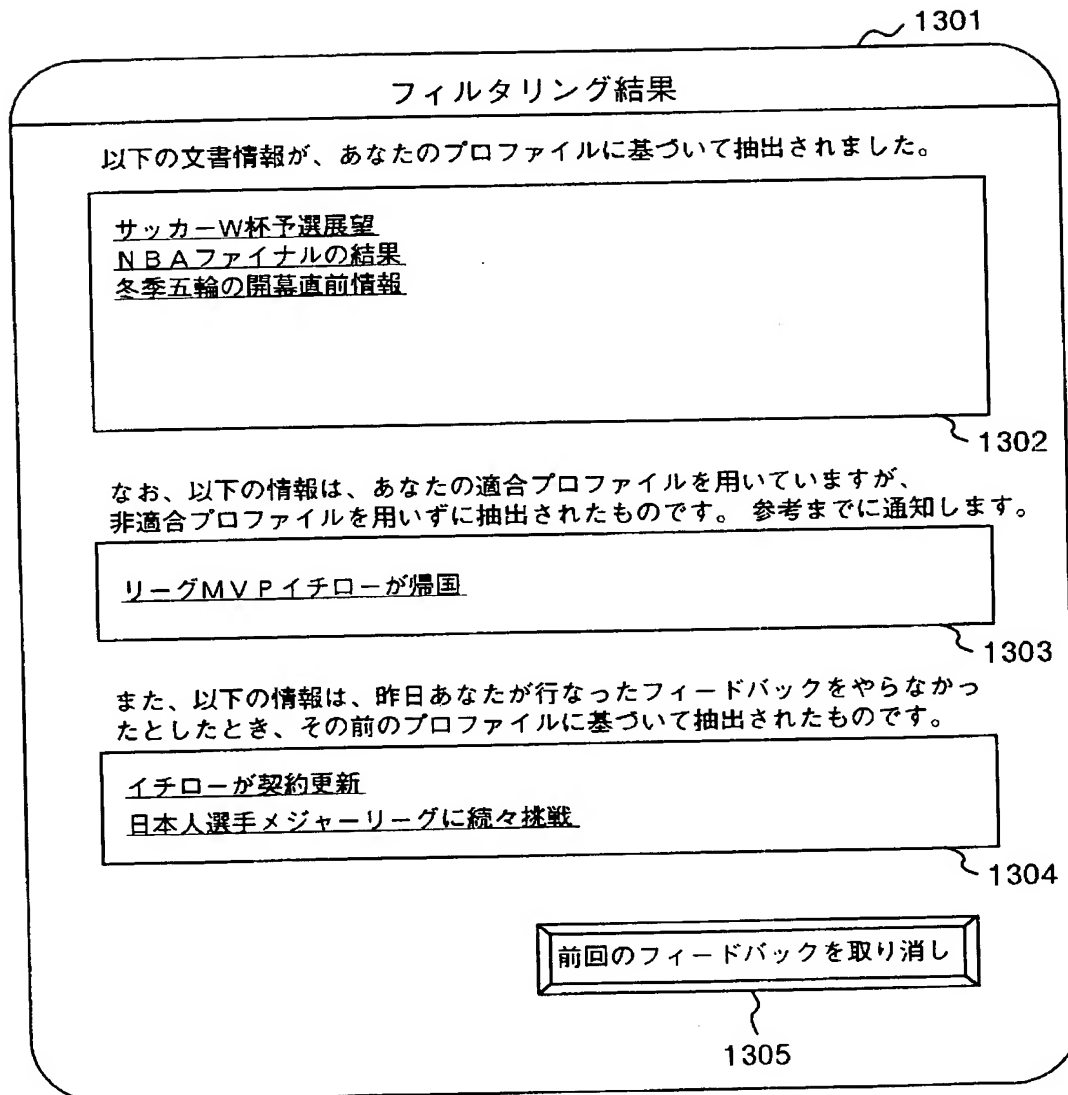
【図 12】

図 12



【図 13】

図 13



【書類名】 要約書

【要約】

【課題】 情報源から送られた文書情報に対してユーザにより設定された配信条件を適用し、この配信条件を満たす文書情報をそのユーザに配信するシステムにおいて、ユーザが配信条件を変更するような操作をしたとき、それによって配信されないことになる文書情報をユーザに提示し、配信条件の変更を評価する。

【解決手段】 文書情報保存プログラム 6 0 2 は、ユーザに配信した文書情報を文書情報格納エリア 6 0 4 に保存する。条件変更リハーサルプログラム 6 0 3 は、配信条件の変更要求を受けたとき、保存された文書情報に変更後の配信条件を適用し、変更後の配信条件を満たさないためにユーザに配信されないことになる文書情報をそのユーザに提示する。

【選択図】 図 6

特願 2 0 0 2 - 2 5 6 8 0 2

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所